DARWIN INFORMATION SYSTEM OF NASA - AN INTRODUCTION

Dennis J. Koga*, David J. Korsmeyer†, and John A. Schreiner‡
NASA Ames Research Center
Moffett Field, CA 94035

Abstract

The DARWIN Information System of NASA is a concept to provide streamlined information access of experimental and numerical test facility data to both internal and external aeronautical customers. This information access consists of broad results from computations and integration of diverse facility instrumentation systems, an efficient information retrieval process, built-in capabilities to present data and results as usable knowledge, and the availability of all the information to customers via remote access connectivity with associated data security constraints. The purpose of this improved information accessibility is to provide aeronautics designers with essential elements to shorten and enhance the efficiency of the design cycle process. The DARWIN Integrated Product Team of NASA Ames was chartered for two years, beginning in March 1995, to develop prototype technology demonstrations for what has become the DARWIN Information System concept. The integrated instrumentation, intelligent database, data visualization, and remote access elements of the overall DARWIN IPT program are introduced and linked together to define the overall system concept.

Introduction

The DARWIN (Developmental Aeronautics Revolutionizing Wind-tunnels with Intelligent systems of NASA) program originates from NASA Ames as an Integrated Product Team (IPT) effort to utilize information systems to augment aeronautical design cycle processes as part of the Ames lead role in the NASA Developmental Aeronautics mission. The DARWIN efforts are built on the previous NASA Ames works of the Integration of Numerical and Experimental Wind Tunnels (IofNEWT) and the Remote Access Wind Tunnel (RAWT) programs, described in detail by Koga, et al.[1]

The NASA Ames' IofNEWT and RAWT projects, in addition to many other NASA and similar efforts, have clearly demonstrated the value of synergistic integration of data from different methodologies, such as experiments and computations, or visualization and quantitative data, or surface and off-body measurements, in order to generate more complete knowledge of the problem at hand. Typically, projects such as these arise from extensive efforts of specialists from individual areas of expertise, joining together for the brief project period. Thus the results produced are typically limited in scope and rather customized applications of data integration.

The intent of the DARWIN Information System concept is to provide an information infrastructure for aeronautics, wherein novel applications of information integration can be easily achieved through simplified access to diverse sources of information. Using Internet web technologies applied to a proprietary network, the DARWIN Information System infrastructure can be thought of as a national NASA intranet for aeronautics. A DARWIN hub will be a central server with links to appropriate customers or sources of information. Utilizing web links instead of a central repository will allow researchers and organizations to control access and integrity of their data and distribute storage requirements of information, but still allow efficient access for authorized data retrieval in a usable format.

The DARWIN IPT was given a two-year charter in March 1995 to provide a prototype of this information technology system for the Ames 12-Foot Pressure Wind Tunnel tests in 1996 and 1997. During the IPT planning process, an aero-centric developmental environment within NASA, capable of providing access to information from any of its testing facilities, was envisioned as the ultimate future goal. One impact that this DARWIN Information System might provide is the

* Assistant Chief, Design Cycle Technologies Branch
† Senior Project Scientist, Member AIAA
‡ Aerospace Engineer

ability [illegible] iterations within a single entry t[illegible] shown in Figure 1, this capabil[illegible]te and efficient access to all rele[illegible]mation from model design and geometry, CFD solutions and grids, wind tunnel and instrumentation data, and ultimately access to rapid prototyping facilities to construct new models in an appropriately short period of a facility test entry. The DARWIN Information System concept is to provide the aeronautical researcher/designer with access to enough relevant information and knowledge in real-time to allow for responsive and informed design decisions.

In order to architect the DARWIN Information System with the desired broad information network capability, the DARWIN IPT is focusing on key implementations and demonstrations of these technology applications in the following product areas: integrated instrumentation suites with test facility data and control systems, intelligent aero-centric capable databases built on meta-data concepts and web retrieval technology, pipelined data visualizers closely coupled to the databases with web viewer processes, and complete remote access to the integrated knowledge sources associated with test processes and databased information.

The aeronautics industry has expressed interest in the DARWIN Information System, and it has been proposed that industry sites establish virtual control rooms or "embassy" sites to connect with the NASA DARWIN system through remote access networks. These embassy site machines can serve as firewalls to maintain corporate computer system and data integrity, yet provide access to appropriate information intended to be accessible by the DARWIN database.

DARWIN Information System Element Overviews

An overview description of the DARWIN IPT product elements and how they work together is provided in the following section. For technical details of these elements, please refer to Schreiner, et al.[2] which describes databasing plans and facility-specific interfaces, and Korsmeyer, et al.[3] which describes web data visualization and remote access methods. Note that except for the customized, facility-specific interfaces, all elements of DARWIN are intended to be easily adaptable to provide information access to a wide variety of sources without imposing data format constraints.

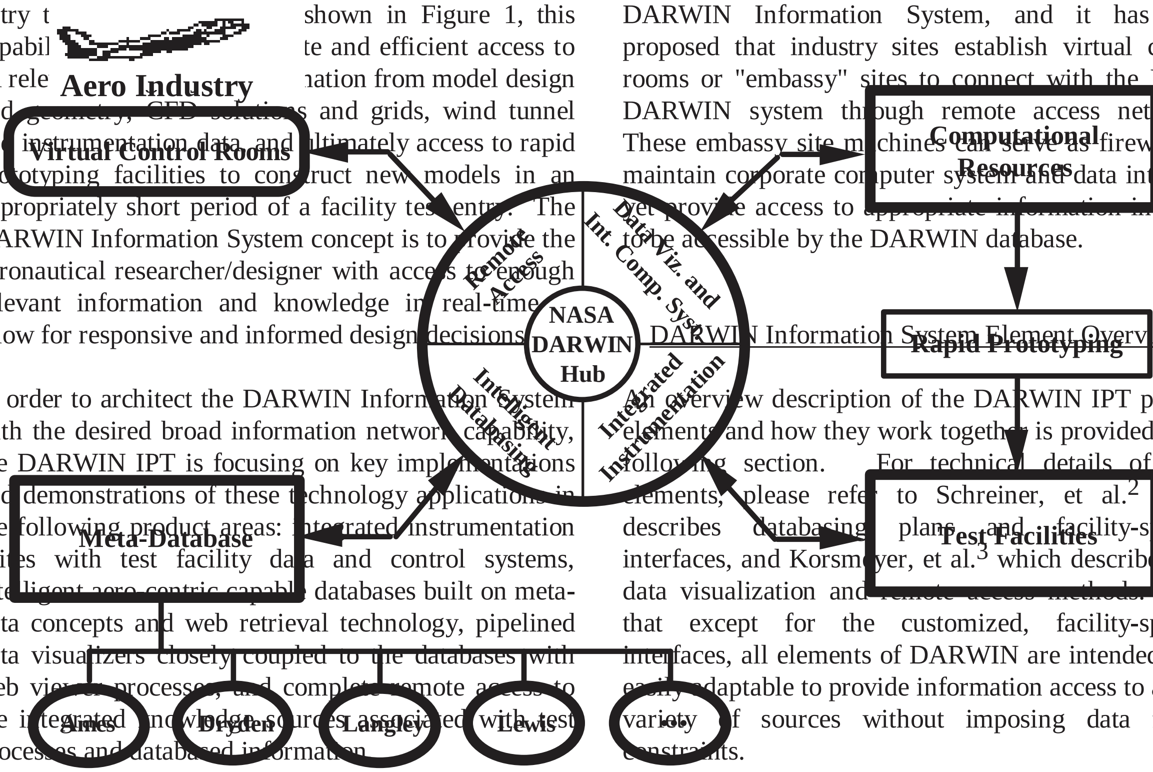

Figure 1. DARWIN Information System Aero-Centric Development Center Concept Schematic

Integrated Instrumentation

The initial step in building the information system is to integrate the individual data sources of a facility (experimental or computational), and also help provide the inter-reliance that many of the new experimental instrumentation systems will have on each other. For example, the Pressure Sensitive Paint (PSP) system has relied on wind tunnel parameter data from the wind tunnel data system in order to properly calibrate results. In the future, to improve accuracy, PSP may also rely on surface temperature measurement systems (such as infrared thermography) and model deformation measurement systems to produce more accurate results. In addition, many of the instrumentation methodologies on the horizon may also rely heavily on information exchange from other acquisition systems such as the wind tunnel data system, thus an intermediate module of information distribution will help prevent overloads which could compromise performance of any particular data acquisition system These are just some examples of how complex individual instrumentation suites will need to work together.

The integrated instrumentation element of DARWIN is called ServIO and is divided into two parts. Note that the DARWIN ServIO process does not take any new data, but rather serves as an information traffic manager. ServIO regulates the synchronous communication between measurement systems and facility control systems, as shown in Figure 2, for maximum test point productivity which is critical since different instrumentation suites will take somewhat different lengths of time to acquire data and often vastly different times to process the data to results. An additional advantage of ServIO synchronization is that future smart scheduling algorithms can be deployed through ServIO to intelligently plan a test matrix wherein the fewer extended instrumentation data processing points could be widely spaced among faster data acquisition points so the entire test process is as efficient as possible.

The ServIO concept was developed to maximize the information flow between data acquisition and instrumentation suites while minimizing the load on any individual system by acting as an information broker in the test facility. In the IofNEWT[1] application, custom software was developed to interface the PSP system with the wind tunnel data system. Adding new instrumentation suites to IofNEWT, with different data requested of the wind tunnel data system, would require new interfaces which would probably contain many redundant pieces of information. As the number of instrumentation suites grows and their complexity increases, it is apparent that a wind tunnel data system, which is usually optimized for expedient collection of a large volume of data, could be adversely affected serving information to numerous systems. In addition, a unique interface needs to be developed for each combination of facility instrumentation system communication in this model.

However, utilizing the ServIO process would minimize the communication load in the integration process by requiring only a single access to any one acquisition system. During the test process, ServIO synchronously collects, organizes, and distributes only the appropriate and needed information to or from each acquisition system. Also in this manner, the data communication and integration concepts developed in a particular facility can be more easily transported to any another facility with different instrumentation by creating a simpler information interface module to ServIO rather than redoing the entire integration effort.

The second role of ServIO, shown in Figure 3, is the asynchronous management of reduced results from instrumentation suites. Since these results may be produced at varying intervals after the test data is taken and the test acquisition process has moved on to subsequent test points, ServIO must properly collect and correlate all information from the various instrumentation used in a particular test point, and then produce the appropriate information set for the DARWIN database. This process is handled by a pre-negotiated file system structure and naming conventions for ServIO data which will be facility and instrumentation technique dependent and thus necessarily customized to each technique as mentioned previously. Again though, once a ServIO-instrumentation interface is developed it should be facility independent and similarly once a ServIO-facility interface is developed, this should be independent of the instrumentation utilized in a test. In order to accomplish this, it is critical that ServIO negotiate the proper information set delivered by each system. This information set is part of the database described next.

Intelligent Database

The DARWIN database is perhaps the most complex issue in terms of overall scope. Whereas the ServIO process has customized facility and instrumentation dependent interfaces, the database must be generally applicable across the vast range of test and computational facilities in order to be truly aero-centric. This is achievable because, for the design cycle relevant

results which will be part of the DARWIN database element, the facility and instrumentation dependent interface issues are handled by ServIO and hidden from subsequent DARWIN processes, except for considerations of the data format of results files. It is useful to note here that the intent of the DARWIN Intelligent Database is to capture result-oriented test information which designers would use to guide a test and perform useful numerical and experimental correlations. Thus, rather than blindly collecting every bit of information associated with test data, or imposing rigorous data format standards (which is often an ill-fated task), the decision was made to create a much more manageable database consisting of intelligent meta-data information, i.e., a subset of data about the data. In this manner, design cycle relevant data could be a direct part of the DARWIN database structure, while the other more supportive data (such as instrumentation calibration information, model geometry files, CFD grid files, etc.) would be indirectly locatable through the DARWIN meta-database, without the burden of duplicating this vast information set within the DARWIN systems.

Surveys of potential users were made to establish the prototype meta-data set. The guidelines for this information were (a) that the meta-data needed to be parameters that users would search across (such as configuration name, Mach number, angle-of-attack, Reynolds number, etc.), (b) that the meta-data and data directly addressable in the DARWIN database would be that information associated with design cycle decisions (pressures, velocity, deformation, etc. and not instrumentation calibration data, etc. which are important but considered as support information) and (c) that the meta-data would be complete enough to contain "pointers" to other filesystems and databases which did have the bulk of the test data. In order to accomplish the last part of the meta-data requirements, a world wide web model of addressing URL's (Uniform Resource Locators) was utilized to provide an aero-centric, virtual capability for the DARWIN databases.

Once the meta-data set was initially defined (by the DARWIN IPT targeting wind tunnel testing and associated computations, but extensible in the future to include propulsion, structural, flight, etc. parameters), the ServIO process was programmed to extract the relevant meta-information from each of the contributing, facility-dependent instrumentation suites and to load the meta-data set into a DARWIN meta-database relevant to the particular test customer. It is important to note that the data security and access control will not be compromised by this seemingly centralized meta-database for the following reasons: (a) there can be independent meta-databases for various customers and/or research programs (although too many independent databases could make desirable cross-configuration comparisons difficult), (b) each customer will have a separate web account with unique account names and passwords to access the database, (c) particularly critical database fields can be further access-controlled, and most importantly, (d) the actual data do not need to reside on DARWIN systems. Clarifying this last point, it is important to remember that the DARWIN database is a meta-database, containing only information about the actual data. While this information is usually somewhat sensitive, the more critical detailed performance data can reside on fully protected (access-controlled, firewall-isolated and encrypted), non-shared resources of which the DARWIN meta-database need only know where to point to retrieve the information using the appropriate URL and possibly subsequent information to decode the files.

Additional advantages of the DARWIN intelligent meta-database concept are that the actual data can be in varying formats and no new standardization need be arbitrarily enforced in order to work with the DARWIN system. While standard file types and structures would be advantageous, it is nearly impossible to achieve consensus for broad ranges of users and data types which DARWIN hopes to serve. However by implementing the meta-database (of which the meta-data itself is an implied standardization of data descriptors, but is also adaptable since it is expandable), the meta-data can contain information about not only where the data resides, but also in what form it can be retrieved. Thus, as with world wide web technology applications, the burden falls on the data retrieval agents and data viewing programs to properly "decode" any data formats (and encryption) utilized by programs which originally produced the data.

Data Retrieval and Visualization Agents

The DARWIN meta-database will access a wealth of design cycle information which will need to be quickly and efficiently accessed in order to truly impact design cycle processes. This will require comprehensive tools to extract usable knowledge for design decisions. While the availability of data will allow future artificial intelligence tools such as feature extraction, trend recognition, etc. to be implemented, the present DARWIN effort is concentrating on more basic, integrated results visualization tools like those used in

IofNEWT to compare experimental and computational results. Coupled in with the standard IofNEWT data plotting comparisons are scripted web interface links to additional visualization tools such as NASA's PLOT3D and FAST (Flow Analysis Software Toolkit) utilities[4,5].

The DARWIN Information System of NASA will be accessed via the proprietary AEROnet network and will provide user interfaces similar to technologies utilized by the Internet (world wide web). A common web browser such as Netscape Navigator™ (which can support Java™ and JavaScript™) will be the only essential user requirement for DARWIN access, with any additional necessary software being downloadable from the DARWIN system after achieving authorized access. The remote access features are described in the next paper section.

Once access has been granted onto the DARWIN Information System, data can be retrieved from the DARWIN database by utilizing widely accepted Standard Query Language (SQL) requests generated from easy to use, DARWIN graphical user interfaces scripts. For example, in the case of a wind tunnel test, a screen displaying current test status and recent run listings will provide the user with summary information, as shown in Figure 4. A point-and-click interface will allow the user to interrogate specific runs by querying appropriate meta-data, then automatically retrieving the actual data, and finally automatically producing data plots and tables, as shown in Figure 5. The user interfaces are customizable for each user to set specific preferences for default plot display formats, and the parameters plotted can be custom-selected from any variable available in the meta-data set. Additionally, the user will be able to apply their own script to process available results in a proprietary manner suited to their own requirements.

A rather unique feature of the DARWIN system is the intimate linkage of data origins provided by the meta-database. The data visualization tools are developed to exploit these traceback capabilities, and features such as clicking on a specific data point of a graph will not only display the actual data value, but also enable the user to further interrogate the source of this "live" data point. For example, a point on a drag polar plot leads to the pressure distribution source plots, then a point on this pressure distribution would lead to the CFD, PSP or pressure tap data sources. In this manner, the user can identify interesting design features and step through the data structures to evaluate the aerodynamic phenomena. The meta-database capability that allows this feature is described by Schreiner, et al.[2] and the web tools to capitalize on this meta-database are detailed by Korsmeyer, et al.[3]. The interconnectivity of the database and the resultant data visualization traceback capability will allow full comparisons of different instrumentation suites, much easier problem diagnoses, and the ultimate utilization of the best, most accurate integrated methodologies available to improve the design cycle process.

Remote Access

The key to building the ultimate DARWIN Information System is aero-centric access to information and the delivery of this information to customers at remote sites. The aerospace industry has estimated that the initial, limited Remote Access Wind Tunnel (RAWT) capability saves at least $25,000 per week in an average wind tunnel test. More complete access to a broader range of design cycle information will far surpass this direct cost savings, in addition to assisting with the faster design of better products which will result from more complete knowledge generated and shared during the design process.

Thus the comprehensive DARWIN implementation of remote access has evolved into much more than the original RAWT application. The improved DARWIN remote access provides extensive capabilities for active participation in tests, a common gateway to aeronautical results, and most importantly, the information infrastructure to implement collaborative engineering, which is a key element for improved design cycle communications and efficiency. In order to achieve this new evolution of remote access capabilities, the following areas have been addressed: networks, access-control, data security, data search tools, collaborative tools, visualization tools, and user access to video, audio, databases, and logbook information. An overview of these concepts follows in this paper, and Korsmeyer, et al.[3] provide complete technical details of the DARWIN remote access element, now referred to as the Distributed Remote Access Machine (DReAM) system.

The DARWIN IPT at NASA Ames has extended the nationwide high speed NAS AEROnet network to all major Ames test facilities, including a new DARWIN data server site and has the capability to expand these subnets to the other NASA aeronautical centers. Access control to this subnetted DARWIN network is handled with a tiered approach including restricted AEROnet user authorization, various levels of

DARWIN subnet filtering, and specific machine and user authentication. In addition, as needed, the remote access process will utilize data security at levels determined by program requirements, including both hardware and software encryption options.

In keeping with the DARWIN Information System concepts, full authorized access to available knowledge will be supported and accessible using the DReAM system tools. As previously mentioned, the key elements for providing this broad information access will be web technology tools, which will allow remote access customers to search information paths as they desire. On-line access to test engineer's logs, control room monitors, database queries, and live instrumentation results are essential elements of remote access. Additional new features will include a graphical user interface control panel for on-demand high resolution video, special remote user camera control and video input stream selection, roaming audio headsets, and enhanced collaboration tools, including new versions of the Silicon Graphics Inc. InPerson™ software and additional new web collaboration tools which are becoming integral components of the latest web browser technology. Since all these features are modular and available through user selectable interfaces, they can be customized for each customer process. The intent of this new version of remote access is to create an environment which will maximize information exchange and also stay current with the latest technology advances in remote collaborative engineering.

Conclusion

The product elements of the DARWIN Information System (integrated instrumentation, intelligent databasing, data visualization tools, and remote access information distribution) are designed to work together to provide the aerospace customer of the future with the necessary information access to greatly improve the design cycle process by gleaning more knowledge from available data. The DARWIN system does not create new data, rather it enables extraction of new knowledge from existing data by providing an information infrastructure and the first broad steps in data integration applications. The ultimate value of DARWIN will be realized as future users develop new knowledge-generating tools utilizing this infrastructure.

The DARWIN program was initiated as a pathfinder for NASA's new Developmental Aeronautics mission. Along its path, the Ames DARWIN IPT has implemented demonstration of the use of high-speed networks and supercomputers in reduction of large data sets for the Microphone Array Phased Processor System (MAPPS) acoustics test in the Ames 40- by 80-foot wind tunnel[6]. In addition, the DARWIN team, in cooperation with NASA Dryden, utilized AEROnet and hardware encryption to provide real-time remote access flight data from NASA Dryden to NASA Langley and an additional industry customer site, in support of the High Speed Research (HSR) program.

The final milestones for the DARWIN IPT are demonstrations of all the integrated DARWIN elements in wind tunnel tests in the Ames 12-Foot Pressure Wind Tunnel. The success of these demonstrations will validate the utility of such an information system, and will be an enabling factor in providing the capability to perform true design cycle iterations in a single entry at a test facility. This capability is envisioned as a valuable goal for maximizing utilization of NASA experimental and computational facilities in the aeronautics industry's design cycle processes. The DARWIN Information System will provide the infrastructure for this capability and thus a true revolution in the way information is delivered to the aerospace customer and how design cycle processes are utilized in the future.

Updated information on the DARWIN, IofNEWT, and RAWT programs can be found on the world wide web at *http://www.darwin.arc.nasa.gov/*.

Acknowledgments

The broad reaching integration efforts of DARWIN rely on a vast variety of contributing personnel from many different areas, and the authors wish to acknowledge such outstanding support and cooperation from the NAS AEROnet group, the NASA Ames CFD and PSP research teams, the Ames wind tunnel facility personnel and the database design groups. The DARWIN IPT is a cross-organizational team bridging the Ames aeronautics and information systems directorates and has received invaluable support from Ames management (in particular Ms. Nancy Bingham, Dr. Walt Brooks, and Mr. Mike George) and programs such as AST and HPCCP. In addition, the efforts to extend the aero-centric DARWIN concepts have been well received and supported by other NASA Aeronautics centers, Headquarters and industry. This interest is well appreciated and should help facilitate future expansion capabilities of the DARWIN Information System for NASA. Finally, the authors would like to thank Kim Washington for her invaluable help in preparing this manuscript.

References

1. Koga, D. J., Schreiner, J. A., Buning, P. G., Gilbaugh, B. L., George, M. W., "Integration of Numerical and Experimental Wind Tunnel (IofNEWT) and Remote Access Wind Tunnel (RAWT) Programs of NASA," AIAA-96-2248, 19th AIAA Advanced Measurement and Ground Testing Technology Conference, New Orleans, LA, June 17-20, 1996.

2. Schreiner, J. A., Trosin, P. J., Pochel, C. A., and Koga, D. J., "DARWIN - Integrated Instrumentation and Intelligent Database Elements," AIAA-96-2251, 19th AIAA Advanced Measurement and Ground Testing Technology Conference, New Orleans, LA, June 17-20, 1996.

3. Korsmeyer, D. J., Walton, J. D., Gilbaugh, B. L., and Koga, D. J., "DARWIN - Remote Access and Data Visualization Elements,", AIAA-96-2250, 19th AIAA Advanced Measurement and Ground Testing Technology Conference, New Orleans, LA, June 17-20, 1996.

4. Walatka, P. P. and Buning, P. G., "PLOT3D User's Manual," NASA TM 101067, NASA Ames Research Center, Moffett Field, CA., March 1990.

5. Walatka, P. P., Clucas, J., Plessel, T. and McCabe, K.R., "FAST User's Guide," NASA Ames Research Center, Moffett Field, CA, November 1994.

6. Watts, M. E., Mosher, M., and Barnes, M., "Microphone Array Phased Processing System (MAPPS)", AIAA-96-1714, 2nd AIAA/CAES Aeroacoustic Conference, May 1996.